Onboard Science Data Analysis: Opportunities, Benefits, and Effects on Mission Design

Paul Stolorz
Jet Propulsion Laboratory
California Institute of Technology

Peter Cheeseman
NASA Ames Research Center

Abstract

Much of the initial focus for spacecraft autonomy has been on developing new software and systems concepts to automate engineering functions of the spacecraft: guidance, navigation and control, fault protection, and resource management. However, the ultimate objectives of NASA missions are science objectives, which implies that we need a new framework for performing science data evaluation and observation planning autonomously onboard spacecraft. We will outline some of the opportunities presented by onboard science processing for transforming the way that scientific data are collected and used by space platforms. The future NASA mission set will feature smaller and more numerous spacecraft in an environment of highly constrained uplink and downlink communications, requiring substantial onboard computation to achieve mission goals. The proposed paradigm will enable mission activities to be directed by scientists without the assistance of a ground sequencing team, robust capture and redirection in making discoveries at the target body, accommodation of the realities of limited communication links, and the return of quality science products from missions.

1 INTRODUCTION: SCIENCE-DIRECTED AUTONOMY

There are several autonomy capabilities that are particularly important in the context of spaceborne science:

- Autonomous identification of features and objects of known interest in onboard acquired images and spectra.
- Prioritization and/or edit of downlink on the basis of reliable recognition of such features and objects.
- Systematic capture of transient science events through integration of autonomous onboard science data processing with autonomous onboard capabilities for retargeting and mission planning.
- Efficient PI-driven redirection of mission activities following scientific discoveries at the target body.

Knowledge on demand is clearly a major ingredient needed to make this vision a reality. A number of intelligent systems technologies can be used to build such an autonomy capability for science. They include data mining technologies, especially pattern recognition, machine learning and knowledge discovery techniques, as well as other capabilities of an autonomous spacecraft, particularly onboard planning. To illustrate the potential for science-driven autonomy, several prototype software systems have been designed by the Machine Learning Systems Group at JPL, and by the Data Understanding Group at NASA Ames Research Center, which tackle well-identified scientific calculations in a spaceborne setting. They include:

- *QuakeFinder*, an automated software package designed to search for sub-pixel surface motions on planetary surfaces.
- *Autonomous Satellite Detection*, onboard software designed to automatically detect and flag natural satellites of small bodies like asteroids.
- *3D Super-resolution,* surface reconstruction software designed to improve image resolution, discover new classes of surface materials and aid data compression.

By performing these identifications onboard a spacecraft in near real-time, these systems are able to provide inputs to autonomous spacecraft executives and planners to enable mission replanning and retargeting of the spacecraft and/or its detectors if a scientific object of great interest and importance is found. In the following sections we outline the design of each of these systems, followed by a discussion of their potential for contributing to science-directed autonomous spacecraft.

2 QUAKEFINDER - SEARCHING FOR SUB-PIXEL SURFACE MOTIONS

A major problem facing scientists in domains such as remote sensing is the fact that important signals about temporal processes are often buried within noisy image streams, requiring the application of systematic statistical inference concepts in order for raw image data to be transformed into scientific understanding.

One class of problems that exploit inference in this way is the measurement of subtle changes in images. Consider for example the case of two images taken before and after an earthquake, at a pixel resolution of say 10 meters. If the earthquake fault motions are only up to 5 or 6 meters in magnitude, a relatively common scenario, then it is essentially impossible to describe and measure the fault motion by simply comparing the two images manually (or even by naive differencing by computer). However, by repeatedly registering different local regions of the two images, a task that is known to be doable to subpixel precision, it is possible to infer the direction and magnitude of ground motion due to the earthquake. The fundamental concept is broadly applicable to many data mining situations in the geosciences and other fields, including earthquake detection, continuous monitoring of crustal dynamics and natural hazards, target identification in noisy images and so on.

One example of such a geoscientific data mining system is QuakeFinder [1], which automatically detects and measures tectonic activity on planetary surfaces by examination of satellite data. QuakeFinder has been used to automatically map the direction and magnitude of ground displacements due to the 1992 Landers earthquake in Southern California. These images were generated by the French SPOT satellite in 1991 and 1992, over a spatial region of several hundred square kilometers, at a resolution of 10 meters, to a (sub-pixel) precision of 1 meter. The system has also been recently applied to the search for tectonic activity on Jupiter's moon Europa, using images obtained over a 17-year interval by JPL's Voyager and Galileo spacecraft.

The system addressed a definite scientific need, as there was previously no area-mapped information about 2D tectonic processes available at this level of detail. In addition to automatically measuring known faults, the system has also enabled a form of automatic knowledge discovery by indicating novel unexplained tectonic activity away from the primary Landers faults that has never before been observed. In the autonomy context, it is potentially a very powerful tool as it can be used to redirect high-resolution imaging devices during the course of a mapping mission when a sufficiently strong signal of surface motion is obtained. For example, future missions to Europa might take along an onboard library of the surface from previous missions, which can be compared on-the-fly with images from the current mission, allowing science PI's or an onboard planning executive to redirect the spacecraft to investigate a dramatic tectonic event at high resolution.

2.1 QuakeFinder algorithm and architecture

The purpose of the basic QuakeFinder algorithm is to detect small systematic differences between a pair of images, which we'll call the "before" image and the "after" image respectively. This is accomplished at sub-pixel resolution by the following method:

1. Match the before and after images by eye as well as possible (i.e. determine the best offsets between the two images in the horizontal and vertical directions).
2. Break the before image up into many overlapping templates, each consisting of, say, 100 x 100 pixels.
3. For each template, measure the correlation between the before template and the after template at the original position determined in step 1), and at the 24 nearest offset positions.
4. Determine the best template offset from the maximum correlation value found in 3).
5. Repeat steps 3) and 4) at successively higher resolution, using bilinear interpolation, or some other interpolation scheme, to generate new templates offset by half a pixel in each direction.

The algorithm relies heavily on the use of sub-pixel image registration for its power. This basic idea is a very useful one that has been successfully applied over the years in a number of fields, especially in the context of image enhancement in undersampled images. Typically, it has been used to automatically account for global effects relating successive images of the same "scene" in an image stream, namely transformations such as translation, rotation and scale changes. We apply the concept here with a highly unusual twist, in that many independent local sub-pixel registrations are performed to disclose the signal of interest, rather than a single global registration.

The QuakeFinder architecture is shown in Figure 1. The first step is application of the basic method to detect the fundamental earthquake motion signal. This step generates a vector field of inferred ground motions from a pair of satellite images. The vector field is then passed to a geometric correction module which automatically corrects for spacecraft artifacts. Upon correction, the resulting displacement map is inspected by geologists for evidence of tectonic activity, with faults being mapped and measured. This information is fed in turn into a further adaptive learning component, in order to refine the fault location and magnitude. This iterative procedure is terminated when sufficient accuracy is obtained. The resulting fault outlines are then registered in a catalog as important events.

2.2 Results for the Landers Earthquake

We obtained the following results from applying QuakeFinder to SPOT data bracketing the Landers earthquake of June 22, 1992. The images are 2050 x 2050 pixels in size covering a 400 square kilometer region of the Southern California desert near the town of Landers. The differences between the two images are extremely subtle and are essentially impossible to detect by eye. Ground motion directions calculated for the Landers quake of June 22, 1992 are shown in Figure 2, superimposed on the 1991 panchromatic SPOT image, in which ground motion direction is encoded in the gray-scale wheel. The major gray-scale discontinuity along the main diagonal of the map shows the position of the fault break inferred automatically by QuakeFinder with no supervised scientific input, based purely on the two raw before and after SPOT images. The black line is ground truth, the known fault location.

Note that the major hue discontinuity corresponds very well to the true fault position, including the bends and steps separating the Emerson fault from the Homestead Valley fault. The general motions are right-lateral, as expected. The motion along the SW block appears to have a north to west trend change as the fault trend itself changes from northerly to more westerly. Thus, the motion tends to parallel the fault, as expected. These observations confirm the value of our approach as an efficient method for automatically detecting and measuring the position of known faults, and surface motion in general.

The QUAKEFINDER System

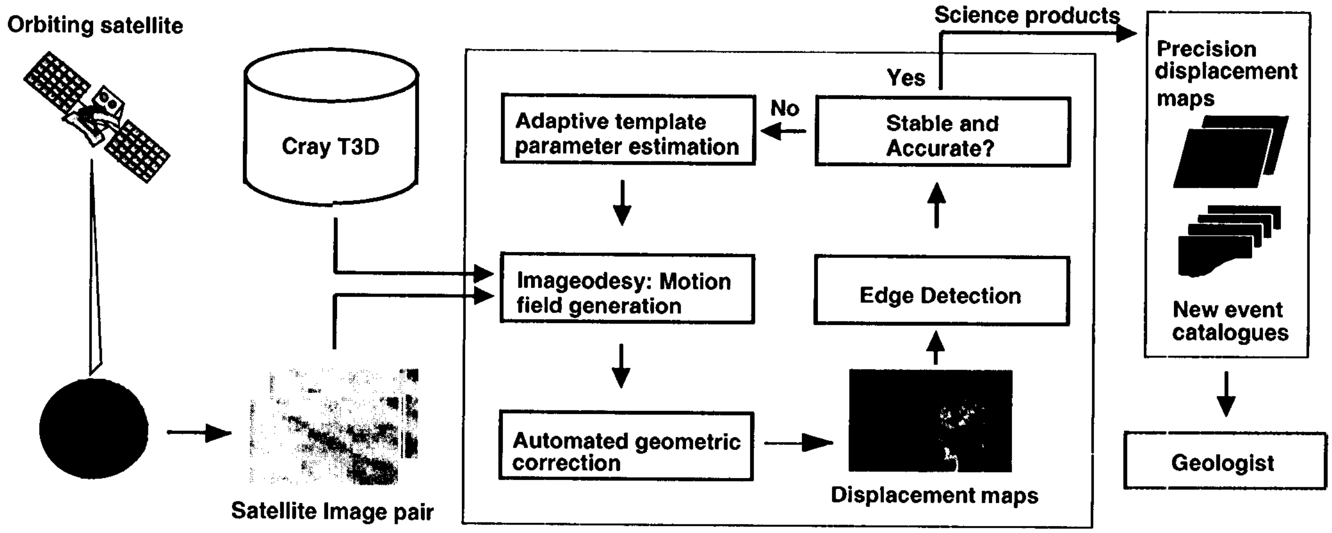

Figure 1. Architecture of the QuakeFinder system.

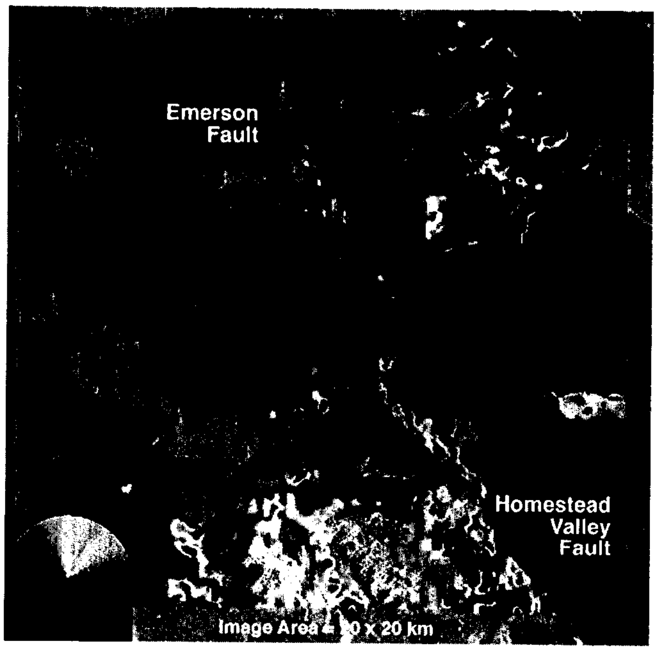

Figure 2. Results of QuakeFinder for the 1992 Landers earthquake

The area-mapped nature of the products generated by QuakeFinder offer an even more interesting capability, namely discovery of entirely new behavior. In the Landers case, it yields suggestive evidence associated with the NE block of the image. This block seems to have two sub-blocks, with relative left-lateral motions between them, suggesting a surface or perhaps sub-surface fault conjugate to the main break. Efforts are currently underway to refine this prediction and to confirm it via field studies. Note that alternative technological approaches cannot easily supply this type of knowledge, if at all. Interferometric SAR can measure small ground displacements in one dimension, along a line perpendicular to the spacecraft trajectory, but cannot supply a full 2D map of motions. Movable seismic detectors located by GPS technology can measure full 3D motion quite precisely, but only at a limited number of individual locations. For these reasons, much of the information displayed in Figure 3 has never before been obtained.

2.3. Implications for Europa and Spacecraft Autonomy

A major effort is currently underway to apply QuakeFinder to search for activity that may have occurred on the ice-covered surface of Jupiter's moon Europa between the visits of the JPL's Voyager spacecraft in 1979 and the recent visit of the Galileo spacecraft [2]. Europa is the smallest of the four Galilean moons of Jupiter. With its icy crust, thin oxygen atmosphere, tidal warming and possible subsurface liquid oceans, Europa is one of the few places in our solar system outside Earth where a legitimate search for life may be conducted. By making detailed measurements of local surface motions and cracks in the crust of Europa, scientists can develop and support theories of its tectonic activity, and in so doing further refine our understanding of the geology and formation of Europa and the solar system.

The surface of Europa is now believed to be substantially younger geologically than once thought, thanks to new high-resolution images obtained in 1996/97 by JPL's Galileo spacecraft. A direct measurement of motion from the time that the Voyager spacecraft visited the planet in 1979 would be an extremely powerful and intriguing result. Also of interest is the possibility of tectonic activity in between flybys of Galileo as the spacecraft enters an extended mission phase in which several flybys of Europa are a core component. Any such activity could be detected, in principle, by the QuakeFinder system.

Quite apart from its value as a ground-based analysis tool, the system could also be implemented on an autonomous spacecraft. For example, it would be possible to equip future missions to Europa with an onboard library of surface images taken by Galileo. These missions could apply the system in realtime during approach or flyby, using initially relatively low resolution images taken by the new mission, to search for regions of possible surface motions since the Galileo mission. Areas that show movement could then immediately by flagged to be retargeted at high resolution as high payoff scientific targets.

There is a clear need in these circumstances for an onboard capability. In a mapping mission, the software provides significant advantages by enabling the prioritization of images for downlink in a highly bandwidth-constrained environment. For any Europa mission, survivability is an issue, for the spacecraft must conduct the mission in the intense radiation environment near Jupiter (the Galileo spacecraft spent only a fraction of its time in the inner Jovian system). Under these circumstances, efficient use of bandwidth to accomplish mission science goals expediently is paramount.

Note in particular that the use of QuakeFinder avoids a common objection that is often raised to the use of science-related autonomy. Important target areas selected by QuakeFinder for high-resolution imaging can be downloaded to earth in their *entirety*, with no compression or corruption of the data, allowing detailed ground-based analysis of the data at all times. The system emphatically does *not* return calculated scientific "results" to earth as its basic product, which would clearly be an activity fraught with pitfalls. Rather, QuakeFinder serves as an intelligent focusing and targeting device, enabling rapid decisions about which data should be collected and downlinked, and with what priority.

With this goal in mind, QuakeFinder has recently been adapted to look at the Europan surface. This problem is substantially more difficult than analysis of the Landers earthquake imagery for several reasons:

1. significant differences in the design and resolution of the two detectors,
2. different imaging geometries due to differing spacecraft trajectories,
3. highly varying illumination conditions.

These difficulties mean that a straightforward automated sub-pixel comparison technique would be prone to substantial systematic errors due to purely geometric and radiometric effects and other detector artifacts. Accordingly, QuakeFinder has been extended with a suite of interactive visualization and analysis modules that allow the basic correlation functions of interest to be studied in detail.

Figure 3 shows a typical image pair of Europa analyzed by the system. The correlation function for a particular inferred offset is also displayed in Figure 4, showing the nature of the correlation surface with respect to different possible offsets of the two images being compared. There are several local minima in this surface which can easily lead to incorrect inferences for the ground motion. A technique has been developed to automatically correct for these problems using prior information on spatial continuity, which successfully locates the correct maximum in the vast majority of cases.

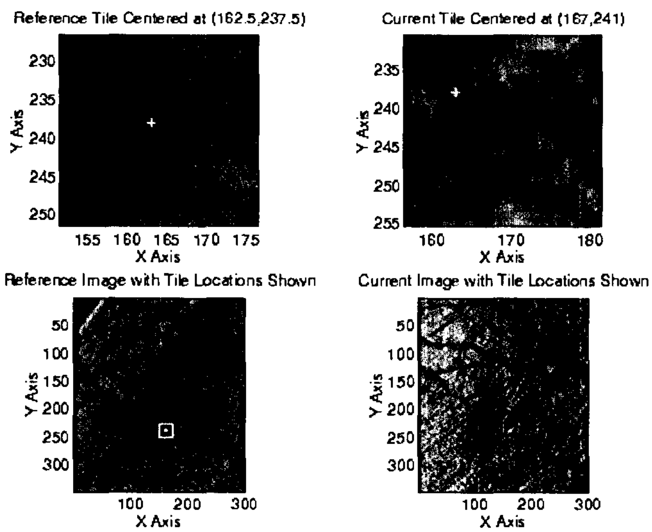

Figure 3. Matched images of Europa taken by Voyager (left) and Galileo (right). The upper images show the two selected regions (squares) magnified to show that they represent the same region.

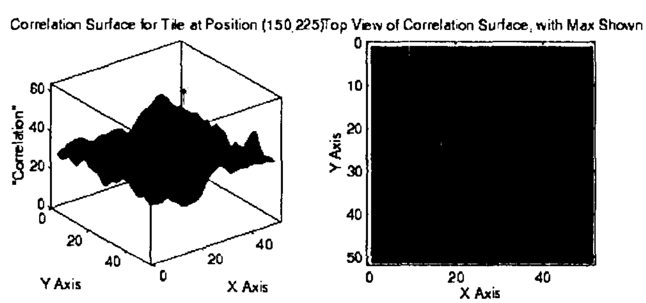

Figure 4. Correlation surface for the image pair shown in Figure 3.

The results display a surprising robustness with respect to these systematic differences for the chosen Europa image pair. The approach reliably reconstructs geometric image transformations using only the raw images as inputs. In the case of the Europa image pair chosen from the Voyager and Galileo image sets, the system automatically corrects for each of these effects. Preliminary results suggest that there is no genuine surface motion over this time frame. This result is not surprising, as the resolution of the images used implies that any surface motion detected must be on the order of hundreds of meters in extent, an unlikely occurrence. There is now an exciting opportunity to use QuakeFinder to search for surface motions in successive passes of Galileo during its extended mission, which will feature several flybys of Europa. These passes will not be subject to large detector differences as the images will be taken with the same device. They will also be at higher resolution than legacy Voyager and Galileo images, enabling searches for surface motion at higher resolution than before.

3. AUTONOMOUS SATELLITE DETECTION

Until relatively recently, the possible existence of natural satellites orbiting asteroids was a controversial notion in the planetary science community [3],[4]. Regarded generally as a relatively rare event, it has not been a prominent consideration in the design of deep space missions. However, the exciting discovery of the satellite Dactyl orbiting asteroid Ida by JPL's Galileo spacecraft has spurred a re-evaluation of the prevailing view about the likely abundance of natural satellites in the solar system.

This new development has ushered in a mindset in which systematic searches for natural satellites can be contemplated as a feasible scientific goal. This perspective suggests several interesting possibilities for scientific discovery driven by spacecraft autonomy. For example, an automated onboard satellite detector offers the potential to detect and flag interesting and unexpected satellites for inspection and retargeting during the course of a mission. We describe here a prototype system that has been designed to perform this task. We plan to extend this initial system to allow the determination of satellite orbits in addition to flagging their existence.

Considerable scientific benefits can be expected to accrue from successful automated detection of satellites onboard autonomous spacecraft. One possibility will be to follow the satellite for sufficient duration to determine its orbit, and from that information to infer the mass of the central asteroid. The size, orbital parameters and composition of a satellite hold significant clues to the origin and age of the asteroid itself. They also have larger implications for understanding the evolution of asteroids as a whole, and consequently understanding their role in the evolution of our solar system. Another possibility is to produce high-resolution images by retargeting the spacecraft or its detectors in order to study important issues such as cratering history, satellite shape and surface geology. In addition, spectral data can be obtained in order to study satellite composition.

The most fundamental task required in any such system is that of identifying candidate satellites in situations in which they consist of a very small number of image pixels (perhaps only one) registering barely above background. This is the situation that exists as a spacecraft first approaches a known asteroid target at far-field, and is the most critical time to flag a new satellite if the spacecraft is to be given time to react to a detection. In these circumstances, satellites cannot be detected from individual images as it is impossible to distinguish them from transient noise sources such as cosmic rays. Several images are required in order to recognize a persistent object. In addition, detector defects and background stars must be removed as sources of noise. A prototype system that performs this task is described in [5].

The most general form of the satellite detection problem is somewhat more complex than this stationary case. It is necessary in general to account for the fact that both the spacecraft and the potential satellite may be in motion over the time period spanned by the relevant image series. A generalized system that achieves this goal for the Ida/Dactyl case has recently been implemented.

The prototype was successfully tested on the Ida-Dactyl images taken by Galileo, for which background stars, detector defects, and cosmic ray hits are the main source of noise. At the farthest range the satellite Dactyl is one pixel large and only a few intensity levels higher than background, and lower in intensity then many of the cosmic ray hits. The detection is performed long before the encounter in order to avoid rapid changes in geometry due to spacecraft motion, and to allow time for non-intrusive processing and potentially, observation replanning.

The software was tested on all the available images of Ida and Dactyl. Its performance was perfect: it successfully detected Dactyl with no false detections. No manual parameter selection was necessary (all the parameters were selected autonomously by built-in procedures). The farthest sequence available was collected at a distance of 171,318 km from the asteroid center, 3 hours 50 minutes before the encounter. Typical Ida / Dactyl images are shown in Figure 5. The successful detection of Dactyl is shown in Figure 6.

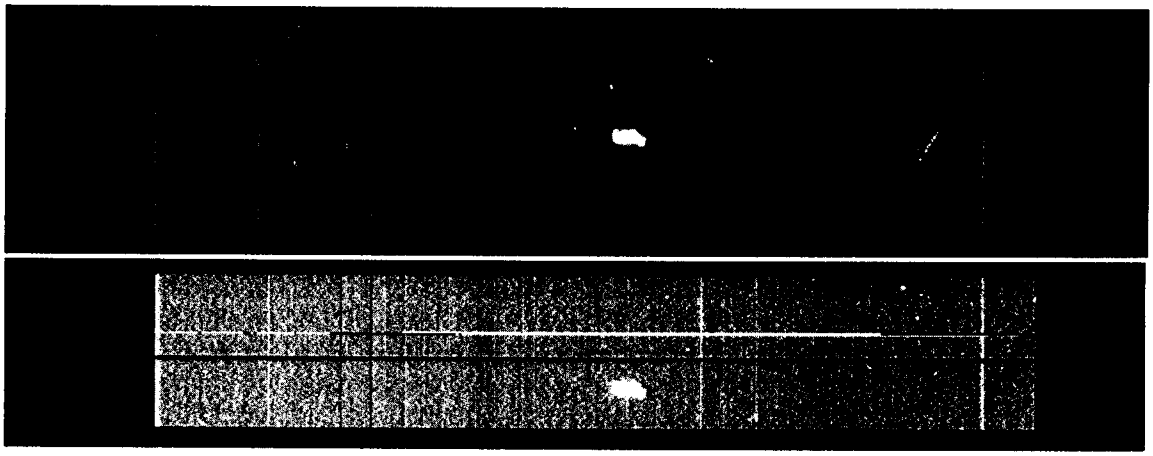

Figure 5. Unprocessed images of Ida and Dactyl.

Figure 6. Detection of Dactyl.

4. ONBOARD SUPER-RESOLUTION

Super-resolution techniques developed at NASA Ames Research Center [6] have proven very successful at improving our knowledge of the fine structural details of static planetary features when several images of the same surface region are available. For example, they have been applied to the analysis of multiple images taken during the 1997 Mars Pathfinder mission. By providing a systematic means for integrating several images taken with different viewing angles, different lighting conditions and different spectral bands, this work provides a mechanism for constructing a surface model at high resolution from low-resolution images. These models have already provided finer feature resolution of the surface than previously possible. They can also be used to capture surface properties such as slope roughness, and mineral content of surface patches. This opens up the possibility of using such models to supply a number of important benefits to autonomous and semi-autonomous missions in the future. Even in rover-class missions where mobility is available, objects at the limit of observability will always remain tantalizing. Super-resolution can bring such objects under closer scrutiny. In general, benefits include:

1. Super-resolution, leading to better onboard decision making in real-time or near real-time.
2. Integration of information. A full 3D surface model provides an excellent summary that can then be put into automated pattern recognition/discovery algorithms.
3. Discovery of new classes of surface materials.
4. Data compression.

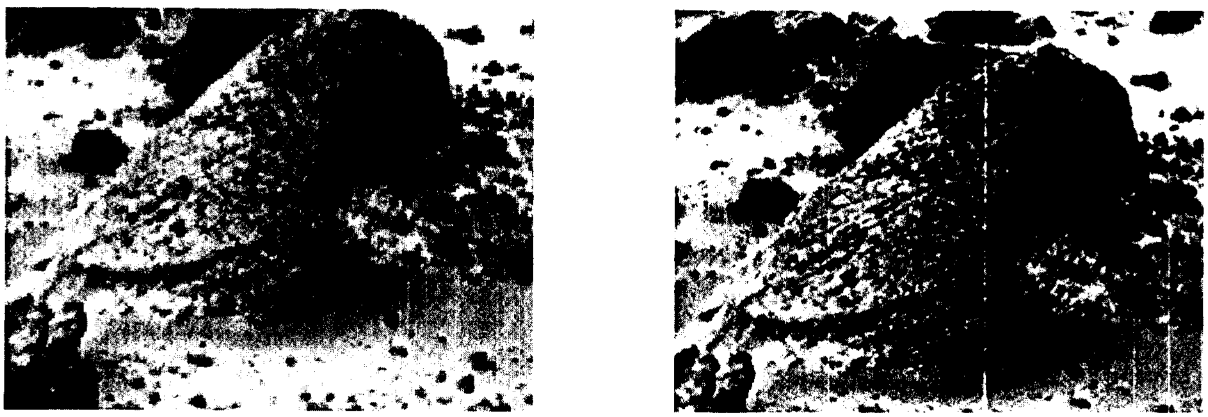

Figure 7. Raw and super-resolved image from Mars Pathfinder mission.

The technique is based upon Bayesian inference, and is described in detail in [6]. The essence of the idea is to invert the standard computer graphics problem of rendering a surface from a given surface model, given specific lighting conditions, etc. Instead, the most likely surface model is inferred from a number of images of the same area. Model parameters include the height, slope and various other surface properties of each surface patch (facet) that make up the area of interest. The approach relies on the development of realistic but computationally tractable surface light scattering models. It also requires an extremely accurate registration between the 3D model and each image. In this respect it relies on the same fundamental registration algorithm as the QuakeFinder work described above. However, the methods differ in their scientific goals, namely greatly improved spatial resolution versus the direct measurement of physical processes to sub-pixel precision.

There are several key technical issues to resolve in the construction of a system like this:

1. Finding very accurate registration between the 3-D model and each image.
2. Finding the right trade-off between the modeling accuracy (i.e. capturing all the known light scattering effects) and computational tractability.
3. Modeling atmospheric effects, particularly clouds (for Earth) and dust (for Mars).
4. Scaling up the algorithms to deal with much larger areas.
5. Identifying the classes found by unsupervised classification techniques on the resulting super-resolved images.

Typical uses of a 3D super-resolution capability include proposed autonomous rover missions to Mars. By providing an onboard image integration capability, the method can provide a 3D surface model suitable for navigation or as input to surface feature recognition software. In addition, it provides an extremely useful tool to planetary scientists and earth science researchers interested in 3D model output, ground cover classification, etc.

5. CONCLUSIONS

A number of future investigations are suggested by the success of the systems described here that will move them even further in the direction of knowledge on demand systems. For example, QuakeFinder can be extended to the continuous domain, measuring very slowly-varying processes instead of abrupt events. This will require the systematic incorporation of scaleable I/O resources to allow the rapid ingestion and processing of continuous image streams. The generality of the basic approach indicates that it will also prove scaleable as detector and satellite resolutions improve. For example, plans are now underway for the development and deployment of satellites with 1 meter resolution or better. Extensions of QuakeFinder will enable physical processes on the scale of centimeters to be straightforwardly detected and measured automatically, opening new avenues of geophysical analysis from spacecraft images.

These techniques show dramatically the power of data mining engines that tackle well-posed scientific problems with a coordinated interdisciplinary approach. There are several other areas that can clearly benefit from the application of data mining techniques such as this, for example global climate change and natural hazard monitoring. One particularly intriguing prospect is the idea of performing monitoring tasks completely autonomously from largely self-directed spacecraft. This is a serious possibility for studies such as plate tectonics, because it is clear that almost no external information is needed to perform the most important geometric corrections.

A number of other scientific goals might be tackled by the science autonomy approach. One possibility is the onboard analysis of asteroidal and planetary craters by automatic means, enabling regions with particularly interesting features to be imaged at high resolution in the case of orbiter and flyby missions, and to be visited in the case of lander missions. Measurement of temporal processes also becomes feasible for a number of situations. These include for example possible sand-dune motion on the surface of Mars, analysis of its ice-cap motions, and studies of Martian storm conditions.

The approach to science-driven autonomy outlined here is specifically designed as the first step in the development of a general robust intelligent onboard science data processing capability. It includes all the characteristics desirable in novel and expensive experiments such as simplicity, robustness, fast processing speed and reliability. The principles underlying the implementation of adaptive science processing tools are generalizable across various missions. There is no doubt that they can dramatically increase the power and scientific value of autonomy concepts in the context of space missions, enhancing their cost-effectiveness by supporting vastly reduced communications bandwidth to earth, while at the same time increasing the value of scientific information returned.

ACKNOWLEDGMENTS

Parts of the research described in this paper were performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

[1] P. Stolorz and C.T. Dean, "QuakeFinder - A Scalable Datamining System for Detecting Earthquakes from Space", *Proc. 2nd International Conf. on Knowledge Discovery and Datamining*, AAAI Press, pp. 208-213, 1996.

[2] P. Stolorz, R. Blom, R. Crippen, J. Roden and R. Granat, "Searches for Sub-pixel Surface Motions on Earth and Europa". *Proceedings of Workshop on Remote Sensing of Planetary Ices*, Flagstaff, AZ, June 1997.

[3] T. Gehrels, J.D. Drummond and N.A. Levenson, "The absence of satellites of asteroids." Icarus 70, 257-263, 1987.

[4] S. Weidenschilling, P. Paolicchi and V. Zappala, Do asteroids have satellites? In "Asteroids II", eds. R. P. Binzel, T. Gehrels, and M. S. Matthews (Tucson: Univ. of Arizona Press), pp. 643-658, 1989.

[5] R. Doyle et al, "Autonomy Enables New Science Missions", STAIF Proceedings, 1997.

[6] P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz and R. Hanson, "Super-resolved surface reconstruction from multiple images", in *Maximum Entropy and Bayesian Methods*, G. R. Heidbreder ed., Kluwer Academic (1996), 293-308.